Rapid Evaluation of Molecular Electrostatic Potential Maps for Amino Acids, Peptides, and Proteins by Empirical Functions

HIDEO NAKAJIMA, OHGI TAKAHASHI, and OSAMU KIKUCHI*

Department of Chemistry, University of Tsukuba, Tsukuba 305, Japan

Received 23 May 1995; accepted 24 July 1995

ABSTRACT

A simple method for evaluating the molecular electrostatic potential (MEP) map without self-consistent field molecular orbital (SCF-MO) calculation is extended, and the parameters for amino acids, peptides, and proteins are determined. In this method, the electrostatic potentials due to electrons in the valence shells are calculated by a set of simple empirical functions at various origins, and those due to the core electrons and nuclei by point charge approximation. For application of the method to amino acids, peptides, and proteins, the functions for the σ and π bonds and lone-pair electrons involved in these species were determined, and the MEP maps calculated by the empirical functions were compared with those calculated by an *ab initio* method. It is shown that the method reproduces correctly the shape of *ab initio* MEP map even for the repulsive MEP region. The method is shown to be very useful for rapid evaluation of reliable MEPs for large biological molecules.

Introduction

The molecular electrostatic potential (MEP) map[1] is very useful to elucidate the molecular interactions in various chemical and biological systems,[2-12] especially the interactions in large biological systems such as drug–receptor interactions[7,8] and enzyme–substrate interactions.[9-12] MEP has been successfully applied to the determination of atomic charges[13] and the molecular similarity index[14] and other molecular properties.[15,16]

The MEP at a point $\boldsymbol{r}$ is given by

$$V(\boldsymbol{r}) = \sum_A \frac{Z_A}{|\boldsymbol{r} - \boldsymbol{r}_A|} - \int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, d\boldsymbol{r}'$$

$$= V_{\text{core}}(\boldsymbol{r}) + V_{\text{elec}}(\boldsymbol{r}) \qquad (1)$$

*Author to whom all correspondence should be addressed.

CCC 0192-8651 / 96 / 070790-16

where Z_A is the nuclear charge of atom A, r_A is the position of atom A, and $\rho(r')$ is the electron-density function at r'. The first term, $V_{core}(r)$, corresponds to the repulsive contribution from the nuclei, and the second term, $V_{elec}(r)$, to the attractive contribution from the electrons. The $\rho(r')$ function can be expressed by the sum of the density functions of occupied SCF-MOs. Since much computation time is required to obtain the SCF-MOs for large molecules, it is difficult to calculate the MEP maps of large biological systems with various conformations. Thus, the method which evaluates MEP without SCF-MOs is required. The point charge models[5,7,9,17] are simple and require less computation time. However, their MEP diverges near the point charge and is reliable only at a long distance from the nuclei in a molecule. Bonaccorsi et al. assumed the transferability of localized MOs (LMOs) from one molecule to another and used LMOs to evaluate the MEPs without MO calculation for each molecule.[18] Although this method gives reliable MEPs, it still requires much computation time. The local density functional approach has been applied to obtain MEPs.[19]

We need a simple model which gives quantitatively correct MEPs with very short computation time and which can be applied easily to large biological systems. This can be accomplished only by using simple functions like those used in the point charge model. We have proposed the empirical functions that evaluate the MEPs without SCF-MO calculation.[20,21] In the method, the MEP is expressed by the sum of the electrostatic potentials (EPs) due to the valence electrons in LMOs and that due to the core atoms. The EP due to electrons in each LMO is calculated as the sum of empirical functions at different origins around the molecule. From previous calculations for the molecules containing the ether type oxygen atoms[20] and for the conjugating systems with π electrons,[21] the method was shown to be useful for the systematic and rapid evaluation of the reliable MEPs of large molecules.

In many enzyme–substrate interactions, the electrostatic interaction is considered to be the most important. Thus in the analysis of such interaction, it is required to know the accurate MEPs of the enzymes and substrates, and it is important to know how the interaction changes the MEP around the active site of the enzyme. To calculate MEPs of such molecular systems rapidly with our method, the parameters which are required for amino acids, peptides, and proteins were determined. They were used to evaluate the MEPs of 11 amino acids and alanylalanine and the MEP change in the active site of the α-chymotrypsin which is caused by proton transfer from Ser 195 to His 57.

Procedure

We used the model localized MO (MLMO) for a valence orbital, which is constructed from the LMO by neglecting the "tail" (the contributions of inner-shell atomic orbitals and of other atoms) and by renormalization. Thus electron-density functions are written as follows:

$$\rho(r) \cong n_i \sum_i^{occ} |\phi_i^{MLMO}(r)|^2 = n_i \sum_i^{occ} \rho_i^{MLMO}(r) \quad (2)$$

where n_i is the occupation number of the ith orbital. The second term of eq. (1) can be rewritten as the sum of the EPs of the MLMOs, $V_i^{MLMO}(r)$.

$$V_{elec}(r) = -n_i \sum_i^{occ} \int \frac{\rho_i^{MLMO}(r')}{|r - r'|} dr' = -n_i \sum_i^{occ} V_i^{MLMO}(r) \quad (3)$$

In this method, $V_i^{MLMO}(r)$ is calculated by using empirical functions at different origins. The inner-shell electrons are not explicitly considered, and Z_A in eq. (1) is replaced by the core charge of the atom A. In this article the method has been extended to amino acids; the functions for the lone-pair orbitals of N and S atoms, the bond orbitals including the N or S atom, and the N π conjugating systems are determined. Figure 1 illustrates the model compounds and their σ, π, and lone-pair fragments, which were used to determine the parameters. It is easy to construct the MLMO with the aid of a minimal basis set, and the STO-5G basis set is appropriate for the present purpose because it has been employed successfully for MEP calculations.[4,6,12] The detailed procedure to determine the functions for the valence electrons has been reported in our earlier articles.[20,21]

Empirical Functions

The EPs of MLMOs of σ and lone-pair orbitals were approximated by three functions ($j = 3$), and

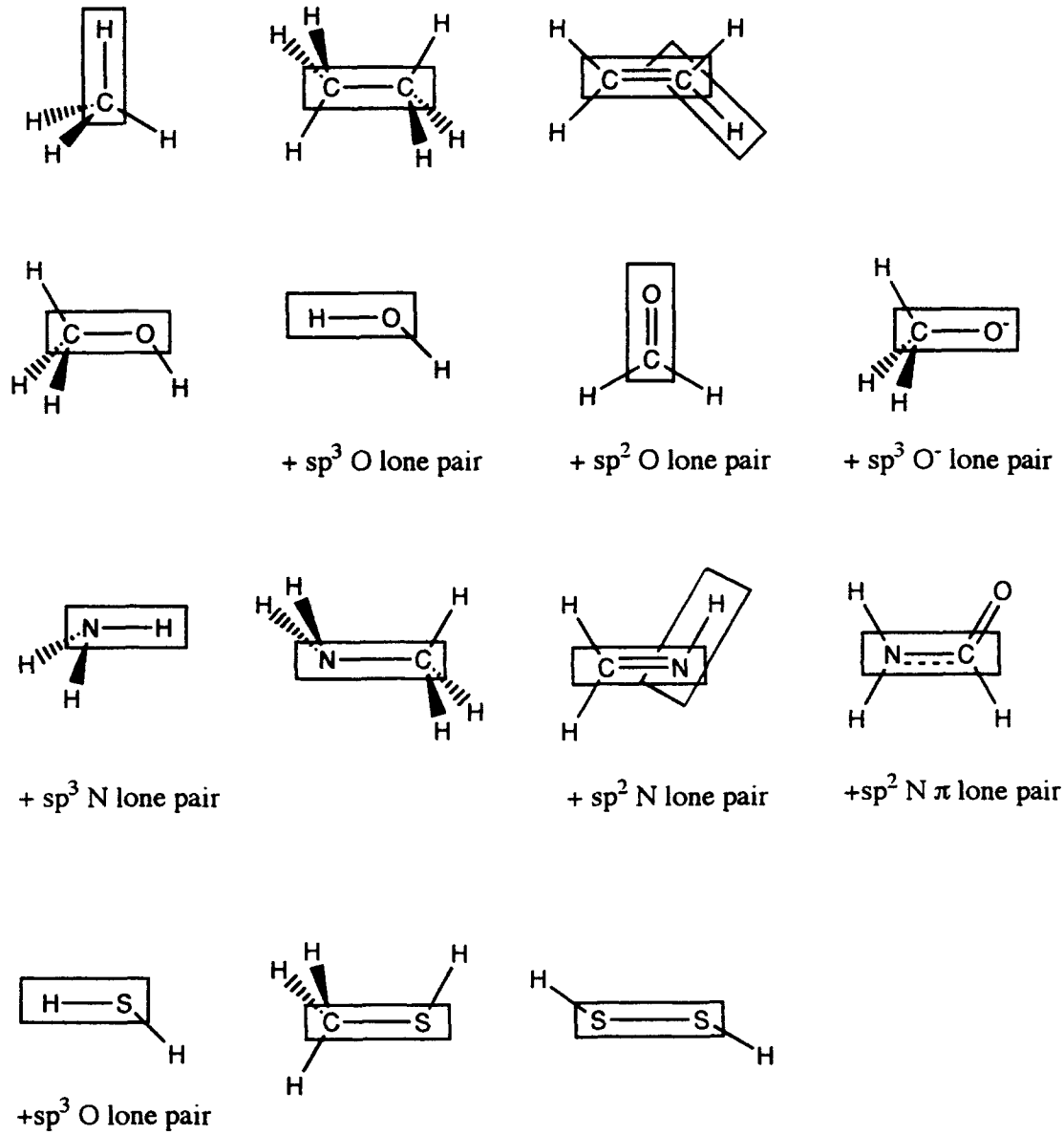


FIGURE 1. Model compounds and fragments used to determine the parameters.

π orbitals by six functions ($j = 6$):

$$F_i(\boldsymbol{r}) = -\sum_j \frac{q_j}{\left[|\boldsymbol{r} - \boldsymbol{r}_j|^2 + a_j \exp\left(-b_j|\boldsymbol{r} - \boldsymbol{r}_j|^2\right)\right]^{1/2}} \tag{4}$$

$$V_{\text{elec}}(\boldsymbol{r}) = n_i \sum_i^{\text{occ}} F_i(\boldsymbol{r}) \tag{5}$$

The origin of the functions of σ and lone-pair orbitals, $\boldsymbol{r}_j$, were located on the symmetry axis of the orbital. The term q_j is the electron population assigned at jth origin. The parameters a_j, b_j, q_j, and $\boldsymbol{r}_j$ were determined by least-squares fitting between the EP of the MLMO and that calculated by the function $F_i(\boldsymbol{r})$, by comparing the EP at 2601 (51×51) points. The parameters for the molecules containing carbon and oxygen atoms have been reported,[20,21] while the parameters which are required for evaluation of the MEP of amino acids, such as the σ and lone-pair orbitals of N and S atoms, were determined here.

C=N π SYSTEM

In the methaneimine type C=N π system, the parameters were determined by the same procedure as the C=C and C=O π systems.[21] In the C=N π systems, the fractional charges, which correspond to the π electron populations, are placed on the symmetry axis of the 2pπ atomic orbitals of the two atoms at the distances of d_C and d_N, and the third electron charge is placed above the center of the C—N bond at the distance of d_{int}. Another set of three charges is placed below the molecular plane symmetrically (Fig. 2). Thus, the six functions are used for the π systems. The amount of the third electron charge q_{int} is the sum of electrons which are transferred from the C and N atoms. Thus, the amounts of the electron charges are

$$q_C = 0.25q_C^{\pi}(1 - \Delta_C) \tag{6}$$

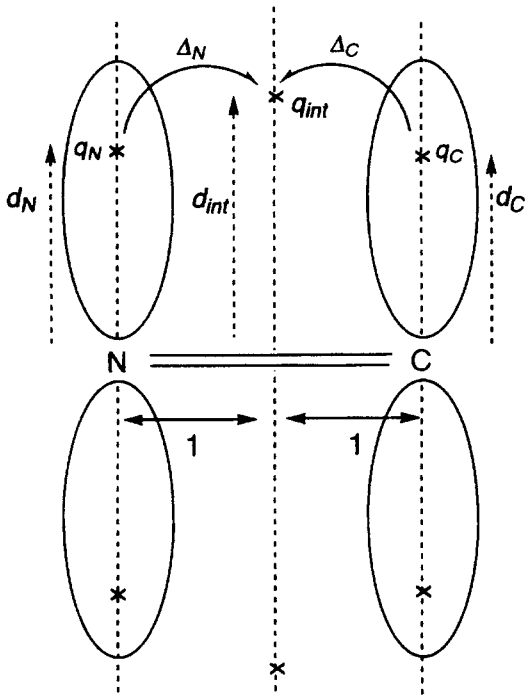


FIGURE 2. Origins of the functions which describe the electrostatic potential due to the C=N π electrons in methaneimine. The q_{int} is placed on the axis which is parallel to the symmetry axes of the two 2pπ atomic orbitals and bisects the C=N bond, Δ_N and Δ_C represent the transfer of the electron, and d_N, d_C, and d_{int} are the distances from C, N and the intermediate point of the C=N bond, respectively. The values of q_C, q_N, and q_{int} are calculated by eqs. (6), (7), and (8), respectively.

$$q_N = 0.25q_N^{\pi}(1 - \Delta_N) \tag{7}$$

$$q_{int} = 0.25(q_C^{\pi}\Delta_C + q_N^{\pi}\Delta_N) \tag{8}$$

where q_C^{π} and q_N^{π} are the π electron population of C and N atoms, respectively, and Δ_C and Δ_N are the ratio of the electron transferred. The π electron population of the C=N bond in a given molecule can be obtained by SCF-MO calculations. However, we set the default valves for the π electron populations which were obtained by the STO-5G calculation of small model molecules in order to avoid SCF calculations for large molecules.

The N—C—O π conjugation cannot be described well by the "C=N + C=O" model. We examined several different models for the formamide type C=N π system and found the model which reproduces the MEP most successfully is that N has the π type lone-pair electrons and q_{int} is transferred only from the C atom (Fig. 3). The charges for the amide N lone-pair orbital are

$$q_{N1} = 0.25q_N^{\pi}\Delta_{N1} \tag{9}$$

$$q_{N2} = 0.25q_N^{\pi}\Delta_{N2} \tag{10}$$

$$q_{N3} = 0.25q_N^{\pi}(1 - \Delta_{N1} - \Delta_{N2}) \tag{11}$$

where q_N^{π} is the π electron population on N, and Δ_{N1}, Δ_{N2} are parameters. The parameter a_j and b_j, origin of the function, r_j, and Δ_{N1}, Δ_{N2} were determined using the N—C—O π system of formamide.

EMPIRICAL FUNCTIONS FOR RING COMPOUNDS

A serious problem found in the construction of MEPs of the aromatic ring compounds is that the functions for σ bonds cannot reproduce well the MEP at the center of the ring. This comes from an accumulation of errors due to the empirical functions for σ orbitals at the center of the ring. We considered the ring correction function for benzene, pyrrole, imidazole, naphthalene, and indole. For benzene, pyrrole, and imidazole, another function whose origin is at the center of the ring was added, and its parameters were determined by comparing the EPs due to the σ orbitals. For example, in benzene, the sum of the EPs of six C—C σ orbitals and six C—H σ orbitals was

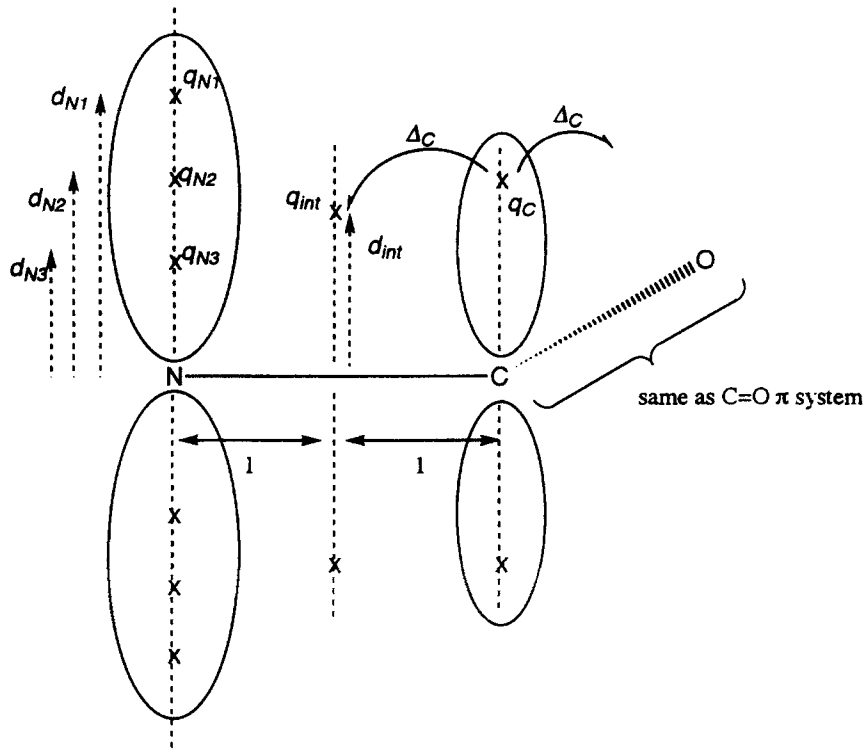


FIGURE 3. Origins of the functions which describe the electrostatic potential due to the N — C — O π electrons in formamide. The six functions are located above and below the N atom. The electron which is transferred only from the C atom is placed on the axis which is parallel to the symmetry axes of the two 2pπ atomic orbitals and bisects the C — N bond. The terms Δ and q have a meaning similar to that in Figure 2.

compared with the sum of the EPs of the empirical functions corresponding to those orbitals and the ring correction function. In naphthalene and indole, three ring functions were used. In the case of the naphthalene ring, two of these are those for benzene, and the third function is located at the center of the condensed C—C bond. In the case of the indole ring, the benzene ring function is placed at the center of the six-membered ring and the pyrrole ring function at the center of the five-membered ring, and the third function is placed at the center of the indole ring.

CORRECTION OF OCCUPATION NUMBER

In our previous studies, electron flow from the lone-pair orbital of the ether oxygen atom or carbonyl oxygen atom to the Cα—X bonds was taken into account by changing the n_i value in eq. (5).[20,21] In the present study, such an electron transfer was considered for the lone-pair orbitals of the sp^3 oxygen, sp^2 oxygen, sp^3 anionic oxygen, sp^3 nitrogen, and sp^3 sulfur atoms. The amount of the transferred charge is determined from the STO-5G Mulliken electron population and is shown in Figure 4. For the lone-pair orbital of the ether-type oxygen atom, for example, 0.01 electrons were transferred from each lone-pair orbital to each methyl group; one methyl group receives 0.02 electrons from the two lone-pair orbitals, and the electrons are assigned to the C—H σ bond (Figs. 4a and 4b). For the carboxyl group, the electron transfer from the carbonyl oxygen atom distorts the MEP map largely and the above correction was not applied to the carboxyl group.

From Mulliken population of the sulfur atom in H_2S, an electron flow is recognized from the sulfur inner atomic orbitals to the valence atomic orbitals of H_2S. To take into account this effect, the core charge of S was set to be 6.08, and n_i of the S—H bond was set to be 2.04.

Results and Discussion

MEP MAPS OF AMINO ACIDS

The parameters required for the MEP calculation of the amino acids are listed in Tables I, II, III,

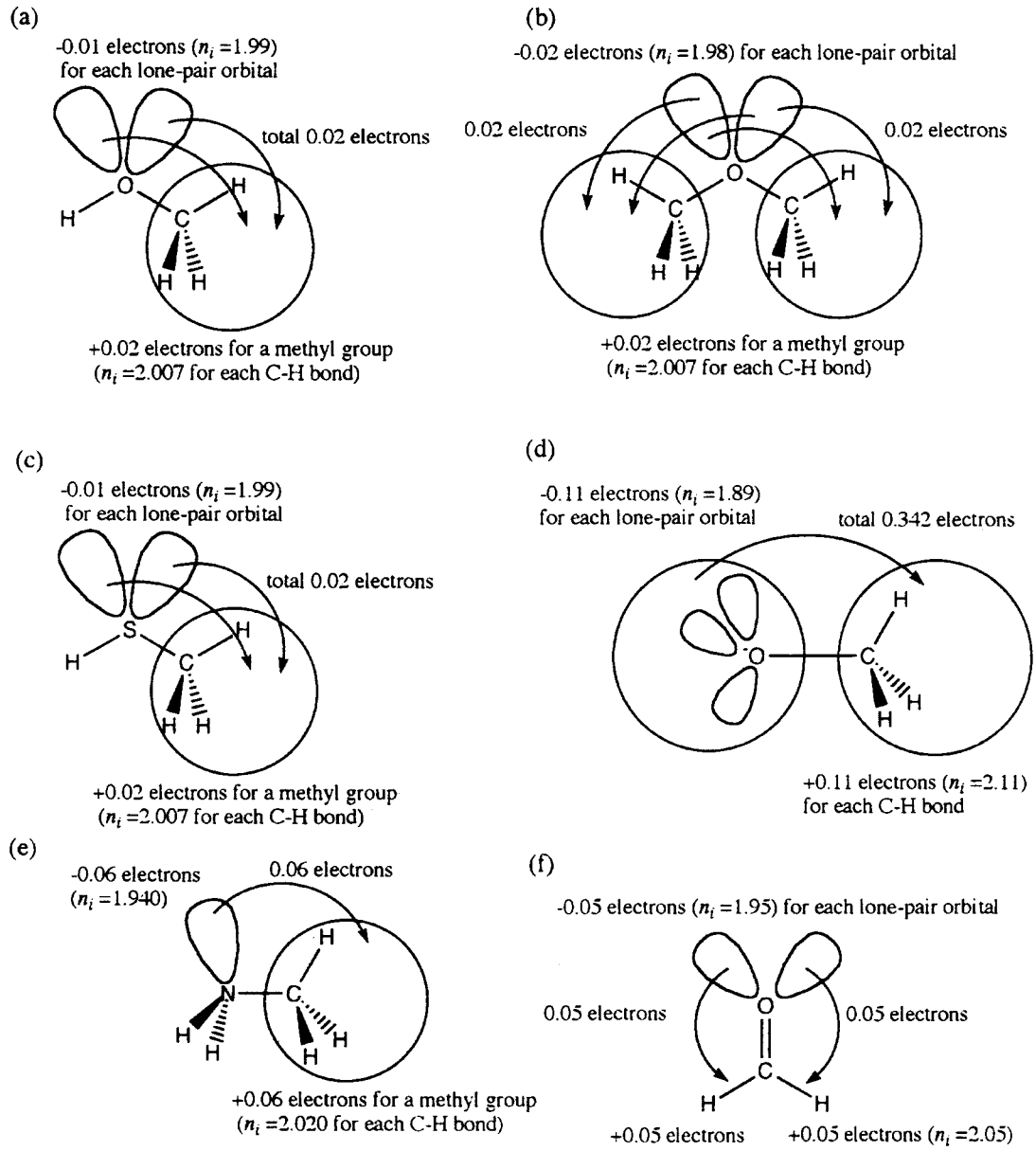


FIGURE 4. The correction of the n_i value for the lone-pair orbitals.

and IV, and the parameters of the ring correction function are listed in Table V. For the π electron population of each amino acid required in the MEP calculation, the values of the corresponding fragment were used. They are shown in Figure 5.

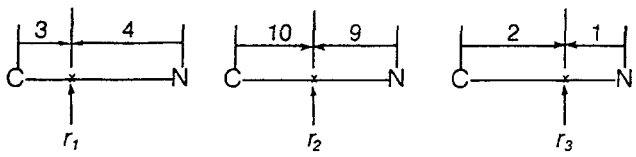
MEP maps of glycine calculated by the empirical function (FUNC MEP) and the STO-5G basis set (STO-5G MEP) are shown in Figure 6. They have three MEP minima: (1) near the nitrogen atom, (2) near the carbonyl oxygen atom, and (3) near the hydroxyl oxygen atom. Their minimum energies are (1) -110.1 kcal mol^{-1}, (2) -71.1 kcal mol^{-1}, and (3) -33.6 kcal mol^{-1} in the FUNC MEP map, and (1) -91.6 kcal mol^{-1}, (2) -58.4 kcal mol^{-1}, and (3) -46.8 kcal mol^{-1} in the STO-5G MEP map, respectively. The function gives more negative MEP for (1) and (2) and less negative for (3), but it correctly reproduces the order of these minimum values and characteristics of the contour lines. In Figure 7, MEP maps of glycine

TABLE I.
Parameters in the Functions for the σ Orbitals.[a]

Parameter	A—B σ orbital							
	$C(sp^3)-H$[b]	C—C[b]	$C(sp^2)-H$[c]	C=C σ[c]	C—O[b]	C=O σ[c]	$C-O^-(sp^3)$	O—H[b]
a_1	0.625	0.661	0.605	0.931	0.828	0.727	0.734	0.259
a_2	0.382	0.418	0.354	0.862	0.333	0.206	0.330	0.355
a_3	0.725	0.661	0.740	0.931	0.144	0.074	0.275	0.742
b_1	0.302	0.393	0.309	0.131	0.563	0.637	0.539	0.453
b_2	0.454	0.523	0.451	0.810	0.513	0.459	0.473	0.664
b_3	0.386	0.393	0.382	0.131	0.450	0.449	0.496	0.467
q_1	0.596	0.751	0.594	0.185	0.675	0.852	0.787	0.026
q_2	−0.253	−0.502	−0.257	0.630	0.303	0.132	0.135	0.373
q_3	0.657	0.751	0.663	0.185	0.022	0.016	0.078	0.601
r_1	8:9	2:3	18:19	3:4	8:9	6:5	14:19	−4:11
r_2	7:4	1:1	19:11	1:1	7:2	13:5	11:4	1:2
r_3	6:1	3:2	11:2	4:3	11:−3	31:−10	4:−1	9:2

Parameter	A—B σ orbital						
	$N(sp^3)-H$	C—N	$N(sp^2)-H$	C=N σ	S—H	C—S	S—S
a_1	0.272	0.660	0.285	1.057	0.786	0.669	1.687
a_2	0.410	0.372	0.389	0.909	1.066	1.149	1.289
a_3	0.846	0.542	0.866	0.633	0.526	0.591	1.687
b_1	0.448	0.403	0.447	0.587	0.395	0.402	0.649
b_2	0.540	0.427	0.559	1.080	0.392	0.489	0.623
b_3	0.455	0.444	0.479	0.180	0.137	0.105	0.649
q_1	0.011	0.762	0.007	0.010	0.049	0.041	0.413
q_2	0.274	−0.403	0.255	0.462	0.738	0.747	0.714
q_3	0.715	0.641	0.738	0.528	0.213	0.212	0.143
r_1	−3:7	3:4	−3:7	18:19	4:11	13:−4	−1:11
r_2	12:23	10:9	10:19	1:1	19:9	54:55	1:1
r_3	11:3	2:1	19:6	19:14	19:−1	3:11	11:−1

[a]The origins of functions r_i are expressed by the ratio which indicates the distances from A and B atoms. In the case of the C—N σ orbital, for example, three r_i are as follows:



[b]From ref. 20.
[c]From ref. 21.

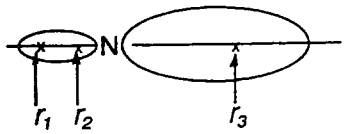
zwitterion are shown. The MEP maps are drastically different from the MEP maps of neutral glycine; protonation at the nitrogen atom makes MEP around the nitrogen atom more positive, and ionization of the oxygen atom makes MEP around the oxygen atoms more negative. The empirical functions reproduce such changes correctly. Figure 8 compares the FUNC MEP values versus the STO-5G MEP values around the (a) neutral and (b) zwitterionic glycine. For the zwitterionic form of glycine, Figure 8b, it is clearly demonstrated that the function gives more negative value in the negative MEP region and more positive value in the positive MEP region.

Figure 9 shows three-dimensional (3D) MEP maps of serine. Isosurfaces were taken at the −10 kcal mol^{-1} level. The shape of the MEP isosurfaces is reproduced well by the function. Four negative regions are seen around the nitrogen atom, the carbonyl oxygen atom, the hydroxyl oxygen atom,

TABLE II.
Parameters in the Functions for the Lone-Pair Orbitals.[a]

Parameter	O(sp^3)[b]	O(sp^2)[c]	O^-(sp^3)	N(sp^3)	N(sp^2)	S
a_1	0.213	0.216	0.213	0.258	0.200	0.477
a_2	0.467	0.461	0.490	0.513	0.519	0.673
a_3	0.485	0.492	0.464	0.593	0.643	1.086
b_1	0.490	0.490	0.571	0.903	0.500	0.251
b_2	0.571	0.572	0.481	0.885	0.637	0.698
b_3	0.610	0.614	0.614	0.597	0.527	0.443
q_1	0.098	0.081	0.106	0.141	0.050	0.226
q_2	−0.876	−0.857	−0.878	−0.688	−0.579	−0.611
q_3	1.778	1.751	1.772	1.547	1.529	1.385
r_1	−0.415	−0.409	−0.416	−0.423	−0.424	−0.410
r_2	−0.006	−0.017	−0.002	−0.021	−0.064	−0.131
r_3	0.207	0.199	0.219	0.260	0.241	0.367

[a]The origins of functions are located on the symmetry axis; r_i is the distance from the nucleus in Å; the electron-rich region is defined as the positive direction.



[b]From ref. 20.
[c]From ref. 21.

TABLE III.
Parameters in the Functions for the π Orbitals.

	A=B π Orbitals[a]		
Parameter	C=C π[b]	C=O π[b,c]	C=N π[d]
a_A	0.567	0.567	0.567
a_B	0.567	0.347	0.459
a_{int}	0.557	0.334	0.446
b_A	0.278	0.278	0.278
b_B	0.278	0.481	0.354
b_{int}	0.258	0.874	0.739
Δ_A	0.138	0.138	0.138
Δ_B	0.138	0.058	0.097
d_A	0.499	0.499	0.499
d_B	0.499	0.389	0.445
d_{int}	0.695	0.621	0.660

[a]The origins of functions are expressed by the distances d_A, d_B, d_{int} from A, B and the intermediate point which divides the A=B bond by 1 : 1, respectively. Δ_A is the ratio of the electron which is transferred from A to the intermediate point (see Fig. 2).
[b]From ref. 21.
[c]The A and B atoms are the carbon and oxygen atoms, respectively.
[d]The A and B atoms are the carbon and nitrogen atoms, respectively.

and the Oγ atom. The negative volume around the Oγ atom overlaps with that around the nitrogen atom. Minimum MEP values of each volume are shown in Table VI. The function reproduces well the minimum MEP value and MEP clouds in the vicinity of Oγ. This is efficient since the MEP near

TABLE IV.
Parameters in the Function for the N — C Group in the N — C — O π System.

Parameter		Parameter	
a_C	0.567	Δ_C	0.138
a_{N1}	0.690	Δ_{N1}	1.003
a_{N2}	0.334	Δ_{N2}	−0.716
a_{N3}	0.433	d_C	0.499
a_{int}	0.446	d_{N1}	0.447
b_C	0.278	d_{N2}	0.255
b_{N1}	0.762	d_{N3}	0.031
b_{N2}	0.495	d_{int}	0.660
b_{N3}	0.642		
b_{int}	0.739		

The origins of functions are expressed by the distances d_C from C, d_{N1}, d_{N2}, and d_{N3} from N, and d_{int} from the intermediate point which divides the C — N bond 1 : 1, respectively. Δ_{N1} and Δ_{N2} are parameters which determine the charge distribution in the N_π lone-pair orbital, and Δ_C is the ratio of the electron which is transferred from C to the intermediate point (see Fig. 3).

TABLE V.
Parameters for the Ring Compounds.

Parameters	Benzene	Pyrrole	Imidazole	Naphthalene	Indole
a	6.053	0.189	1.542	4.402	−0.796
b	0.121	1.004	0.106	0.428	0.304
q	−0.076	−0.066	−0.304	0.029	−0.014

The function is placed at the center of the ring (see text).

the side chain is sometimes more important than that of the main chain in the interaction of proteins.

The 3D MEP maps of aspartate anion are shown in Figure 10. Its amino group was treated as a protonated form, while two carboxyl groups were treated as anionic forms. Isosurfaces were taken at the -100 kcal mol^{-1} level. The map was obtained by using the same parameters as the neutral molecules for the carboxyl group except the π electron population, which is given for the carboxylate anion in Figure 5. The MEP minimum values are -221 kcal mol^{-1} in the main chain region and -230 kcal mol^{-1} in the side chain region in the FUNC MEP map, whose STO-5G values are -198 kcal mol^{-1} and -201 kcal mol^{-1}, respectively. It is shown that the empirical function and parameters reproduce well the MEP clouds of ionized fragments if the corresponding π electron population is used.

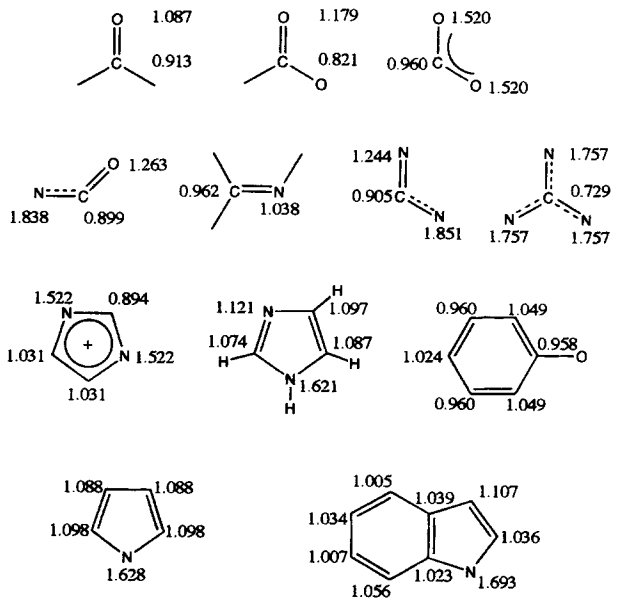


FIGURE 5. The π electron populations of the model compounds. These values were used in this study for the fragments in amino acids and peptides.

Comparison of 3D MEPs of 11 amino acids—glycine, alanine, serine, cysteine, valine, aspartate, asparagine, methionine, histidine, tyrosine, and tryptophan—revealed the following characteristics for the FUNC MEPs of amino acids:

1. In the main chain region, the potential well of the FUNC MEP is deeper than that of the STO-5G MEP by about 20 kcal mol^{-1} near the nitrogen atom, and by about 10 kcal mol^{-1} near the carbonyl oxygen atom, while it is shallower by about 5 kcal mol^{-1} near the hydroxyl oxygen atom (Table VI). However, the order of their potential energies, N > C=O > OH, is reproduced correctly for all amino acids.
2. For the side chain region, the differences of minimum energies between FUNC MEP and STO-5G MEP are less than 10 kcal mol^{-1} (Table VI); the function reproduced the STO-5G MEP better than the main chain region. The empirical functions are useful, especially in calculating MEPs of proteins, such as the enzyme active site, where the functional groups of the residues play an important role in the catalytic process and in interactions of proteins.
3. For the negative MEP region, the averaged FUNC MEP value is more negative than the averaged STO-5G MEP value by 1 ~ 3 kcal mol^{-1}, and, for the positive MEP region, the averaged FUNC MEP value is more positive by 1 ~ 2 kcal mol^{-1}.
4. The correlation coefficient between the FUNC MEP and the STO-5G MEP, r, ranges from 0.93 to 0.97, indicating that the overall MEP shape is well reproduced by the functions.

For zwitterionic amino acids and those having an ionized side chain, such as the aspartate anion and the histidine cation, the function reproduces the negative region more negative and the positive region more positive (Figs. 7 and 8b). But the

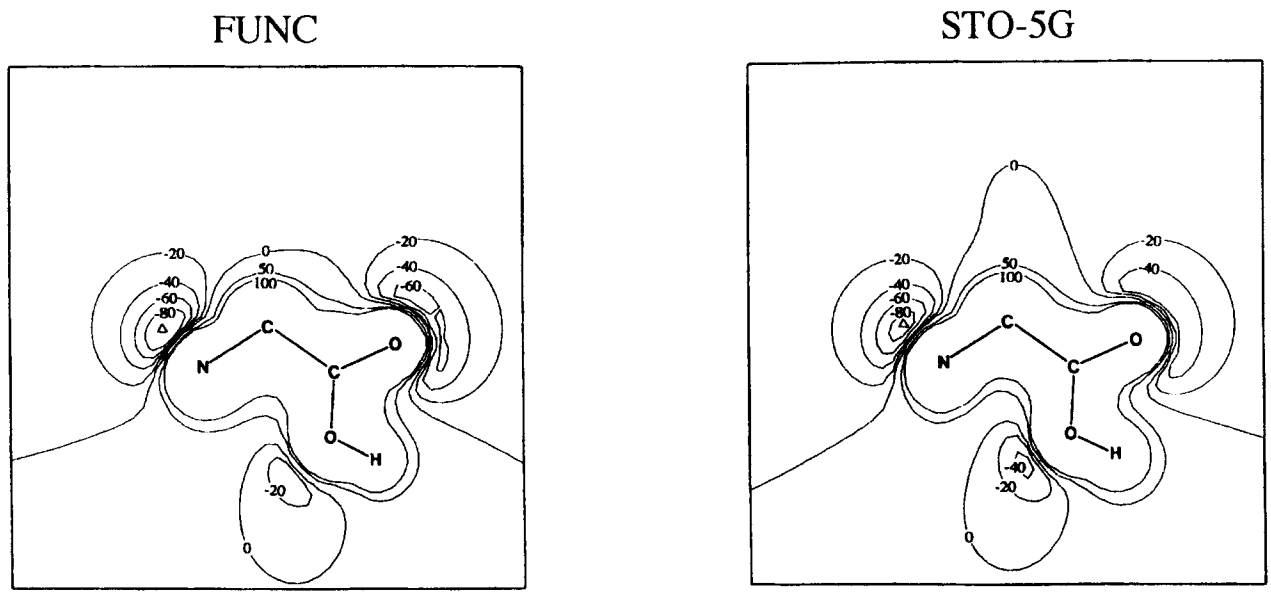


FIGURE 6. FUNC MEP map and STO-5G MEP map of glycine calculated on the plane which includes the N — Cα — COOH group. Values are in kcal mol^{-1}. The potential minimum is shown by a triangle mark, and the minimum value is −110.1 in FUNC and −91.6 in STO-5G.

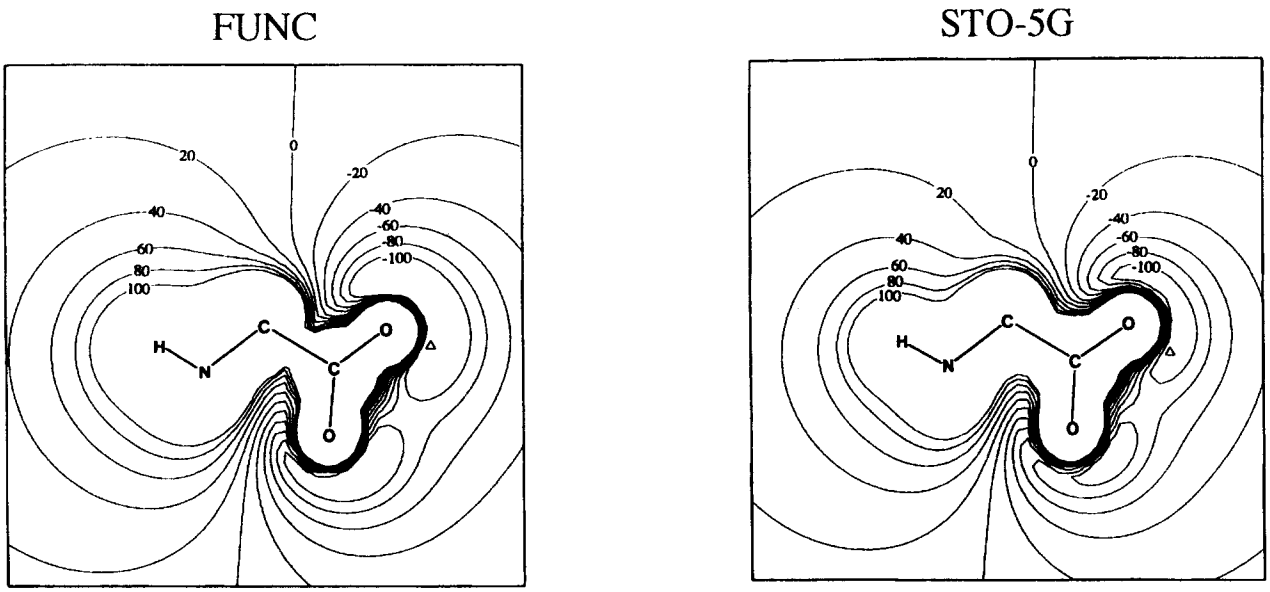


FIGURE 7. FUNC MEP map and STO-5G MEP map of glycine zwitterion calculated on the plane which includes the HN — Cα — COO group. Values are in kcal mol^{-1}. The potential minimum is shown by a triangle mark, and the minimum value is −165.1 in FUNC and −126.7 in STO-5G.

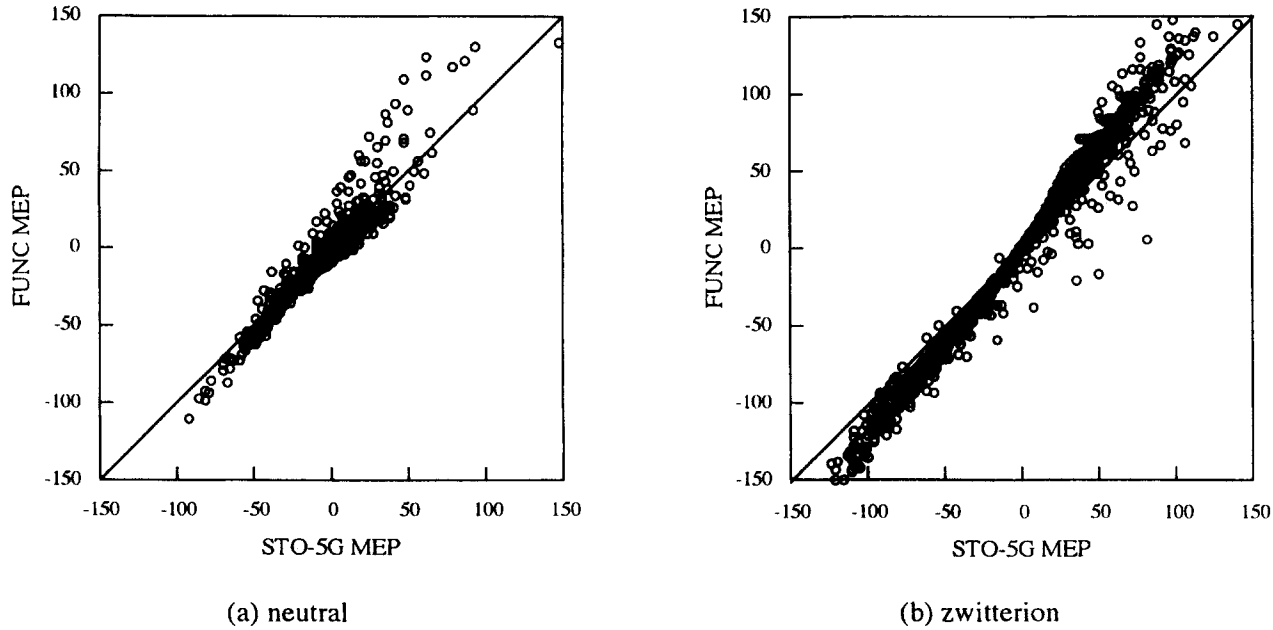


FIGURE 8. Correlation between the FUNC MEP values and the STO-5G MEP values for (a) glycine and (b) glycine zwitterion. The comparison was made for the positions which were located on the plane which includes the N — Cα — COO group and were at more than 1 Å distance from any atoms.

correlation coefficient is 0.99 for all amino acids. In calculation of the ionic groups, such as the carboxyl group and the imidazole cation, the function parameters determined for those in neutral form were used except the π electron population, which is shown in Figure 5. It is confirmed that the empirical functions and the parameters are successfully applicable to the systems in which the electron population changes.

MEP MAP OF ALANYLALANINE

Figure 11 shows the 3D MEP maps of alanylalanine in the neutral form (Scheme 1). Isosurfaces were taken at the -10 kcal mol^{-1} level. The function reproduces the overall shape of the STO-5G MEP map well. Four negative regions are seen, and their minimum energies are listed in Table VII. Three of them correspond to the negative

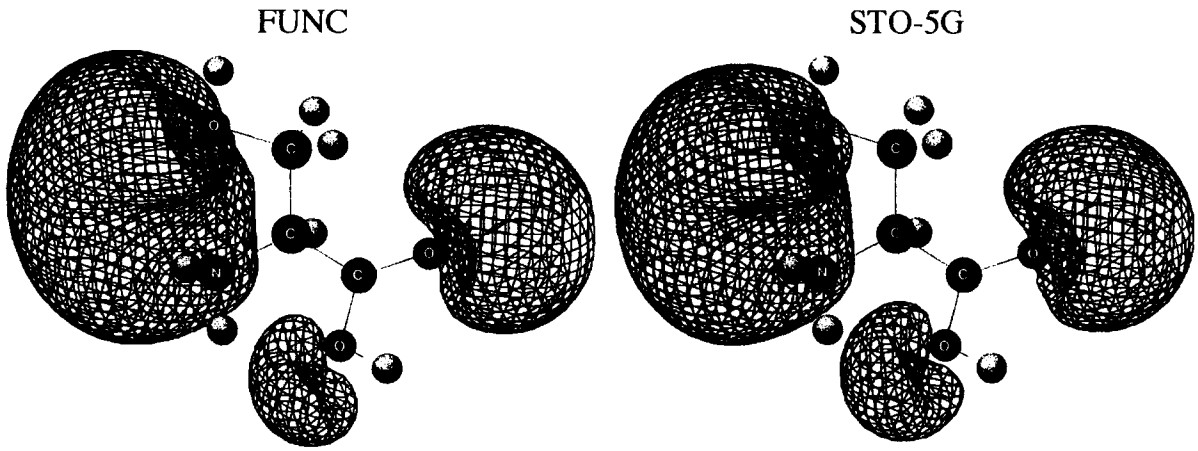


FIGURE 9. 3D maps of FUNC MEP and STO-5G MEP of serine. Isosurfaces were taken at the -10 kcal mol^{-1} level. The minimum values of each clouds are listed in Table VI.

TABLE VI.
Comparison of Minimum Energies of the 3D MEP of 11 Amino Acids.

Molecule		N	C=O	OH	Side chain			
Glycine	FUNC	−110.1	−71.1	−41.1				
	STO-5G	−91.6	−58.3	−46.8				
Alanine	FUNC	−120.0	−71.4	−42.2				
	STO-5G	−92.8	−58.9	−47.4				
Serine	FUNC	−119.5	−67.7	−41.8	OH	−70.2		
	STO-5G	−99.5	−56.6	−47.8		−67.3		
Cysteine	FUNC	−116.6	−64.5	−42.1	SH	−38.0		
	STO-5G	−92.0	−50.7	−47.4		−31.6		
Valine	FUNC	−114.8	−67.0	−42.1				
	STO-5G	−94.7	−58.5	−47.1				
Aspartate	FUNC	−103.8	−57.3	−40.4	C=O	−47.3	OH	−58.1
	STO-5G	−91.3	−48.6	−48.1		−45.5		−53.1
Asparagine	FUNC	−106.5	−57.0	−36.8	C=O	−80.7	NH_2	−26.8
	STO-5G	−86.4	−51.2	−42.2		−71.5		−32.8
Methionine	FUNC	−108.1	−66.6	−39.5	S	−32.0		
	STO-5G	−91.0	−52.3	−46.2		−28.8		
Histidine	FUNC	−99.4	−61.2	−37.8	N	−106.6		
	STO-5G	−83.8	−51.4	−41.3		−101.0		
Tyrosine	FUNC	−111.3	−66.7	−40.7	OH	−66.8		
	STO-5G	−90.2	−58.8	−47.8		−60.8		
Tryptophan	FUNC	−116.1	−58.7	−41.2	indole ring		−21.0	
	STO-5G	−92.9	−52.7	−45.7			−18.5	

Values are in kcal mol^{-1}.

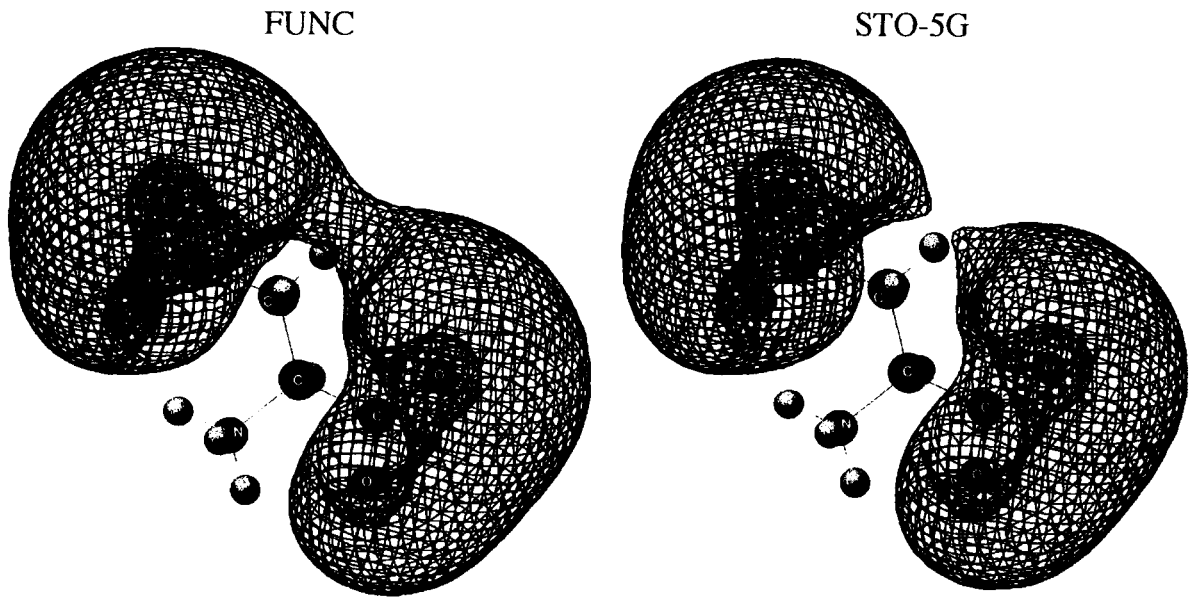


FIGURE 10. 3D maps of FUNC MEP and STO-5G MEP of aspartate zwitterion. Isosurfaces were taken at the −100 kcal mol^{-1} level.

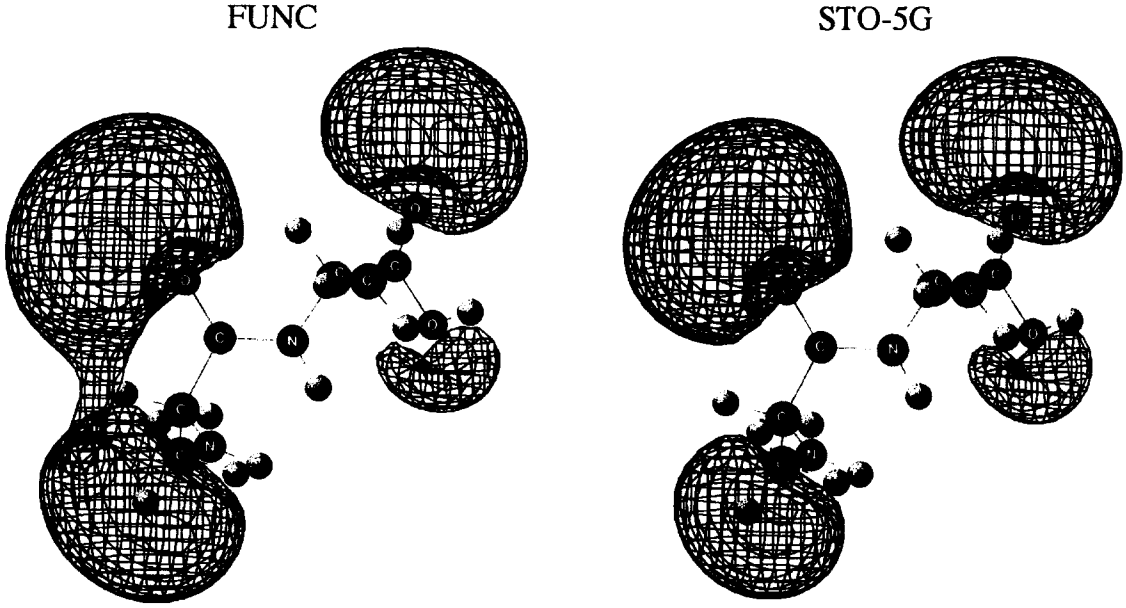
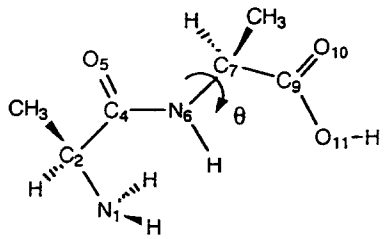


FIGURE 11. 3D maps of FUNC MEP and STO-5G MEP of alanylalanine ($\theta = 0°$). Isosurfaces were taken at the -10 kcal mol^{-1} level. The minimum values of each clouds are listed in Table VII.

regions observed in the 3D MEP map of alanine. The fourth is seen around the peptide oxygen atom. These negative regions are characteristic of the MEP map for the protein main chain, and their relative volumes and relative minimum MEP values are calculated correctly by the empirical functions. Peptides change their conformation by rotation around the N—Cα bond. The MEP change caused by this conformational change was examined. Figure 12 shows the MEP for the conformation which is obtained by the 90° rotation of the N_6—C_7 bond ($\theta = 90°$). As may be seen from Figures 11 and 12, the empirical functions describe quantitatively the change in MEP caused by the conformational change.



SCHEME 1. Alanylalanine.

MEP MAP OF THE ACTIVE SITE OF α-CHYMOTRYPSIN

With the empirical functions determined in this study, we calculated 3D MEP map of the active site of α-chymotrypsin. It catalyzes the rupture of a peptide bond caused by a nucleophilic attack of ionized Oγ of Ser 195.[22] Its active site is mainly composed of the catalytic triad, His 57, Asp 102, and Ser 195. Ionization of Oγ of Ser 195 is accomplished by a proton transfer from Ser 195 to His 57. As the active site, we choose 22 amino acids which are located within 7 Å from the active serine (Ser 195).[11] The crystal structure of α-chymotrypsin,[23] obtained from the Brookhaven Protein Data Bank, was used for the coordinates of these 22 amino acids, and hydrogen atoms were properly added. The NH and CO groups which were artificially created at the end of the selected amino acids were converted into the NH_2 and COH groups, respectively. Asp 102 and Asp 194 were treated in anionic forms. As a whole entity, the enzyme active

TABLE VII.
Comparison of Minimum Energies of the 3D MEP of Alanylalanine.

θ		N_1	O_{10}	O_{11}	O_5
0°	FUNC	−117.0	−67.2	−40.4	−69.6
	STO-5G	−91.1	−57.0	−43.3	−62.8
90°	FUNC	−116.1	−72.4	−51.8	−68.4
	STO-5G	−91.5	−60.6	−54.7	−63.3

Values are in kcal mol^{-1}.

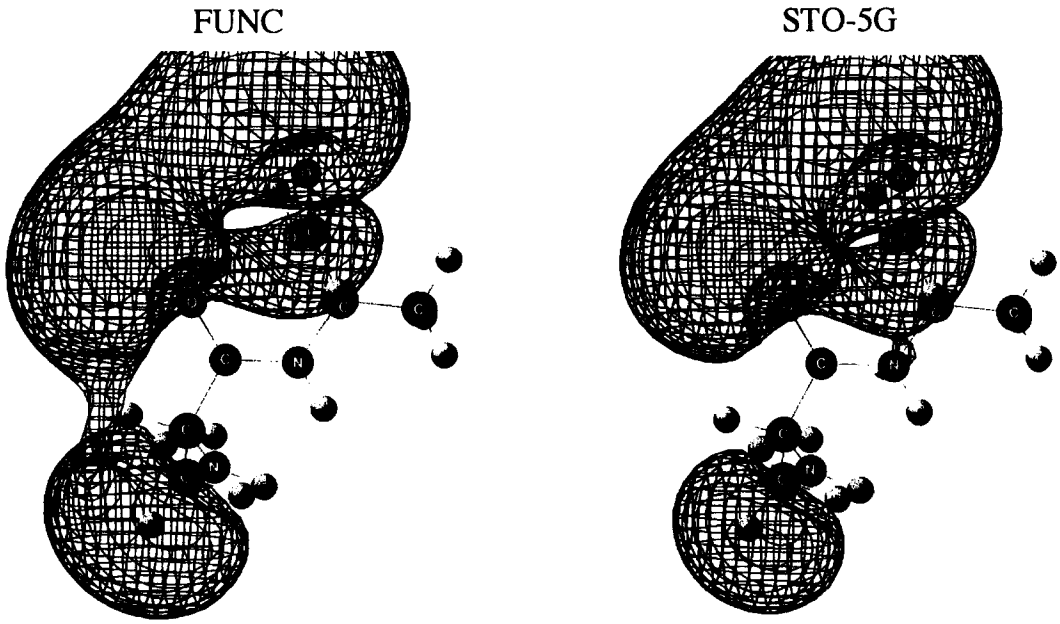


FIGURE 12. 3D maps of FUNC MEP and STO-5G MEP of alanylalanine ($\theta = 90°$). Isosurfaces were taken at the -10 kcal mol^{-1} level. The minimum values of each clouds are listed in Table VII.

site is electrically neutral because of the counterions. Since their positions are not known, the charges of the functions [q_j in eq. (4)] of all points are decreased proportionally so that the total charge is kept at 0. Figure 13 shows the 3D MEP maps of the active site (a) before and (b) after the proton transfer occurs. Isosurfaces were taken at the -80 kcal mol^{-1} level (red) and the -30 kcal mol^{-1} level (blue). Before the proton transfer, a large MEP cloud so-called nucleophilic suction pump,[11] which spread over Oγ of Ser 195, C=O and Oγ of Ser 214, Oδ1 and Oδ2 of Asp 102, the main chain region of Ala 56, and C=O of His 57 is observed. The MEP clearly shows the movement of the negative cloud from Asp 102 to Ser 195 with the proton transfer from Ser 195 to His 57 (Figs. 13a and 13b), and it shows the increase of the nucleophilicity of Oγ of Ser 195 successfully. Large and deep MEP clouds are generated around the ionized Asp 102 even after the proton transfer. The negative potential in this region plays an important role in stabilization of His 57 cation generated by the proton transfer from Ser 195,[24] and the FUNC MEP map shows that effect clearly.

COMPUTATION TIME

Table VIII shows the computation time required for the MEP calculation for 132,651 (51 × 51 × 51) points around the molecule. Calculation was carried out on the HP 755 workstation. In the case of tryptophan, for example, the computation time of the FUNC MEP is 1/1000 of that of the STO-5G MEP, and the use of the empirical functions becomes more efficient for larger molecules; for the case of the active site of α-chymotrypsin, the computation time would be 1/200,000. The empirical functions calculate the MEP maps for a large molecule, such as an enzyme, with a very short computation time. They are very useful in analyzing the molecular interactions in biological systems where evaluation of the MEP maps is required for various conformations.

Conclusion

The method for rapid evaluation of the molecular electrostatic potential map by empirical functions has been extended to amino acids, peptides, and proteins. The parameters for systems which contain nitrogen and sulfur atoms and the additional parameters for the five ring compounds have been determined using the model compounds and comparing the STO-5G MEPs and those obtained by the functions. The parameters reproduced well the MEP for the amino acids and peptides. Although the absolute values for the potential wells are larger than the STO-5G values, the overall shapes of the MEP maps, the relative

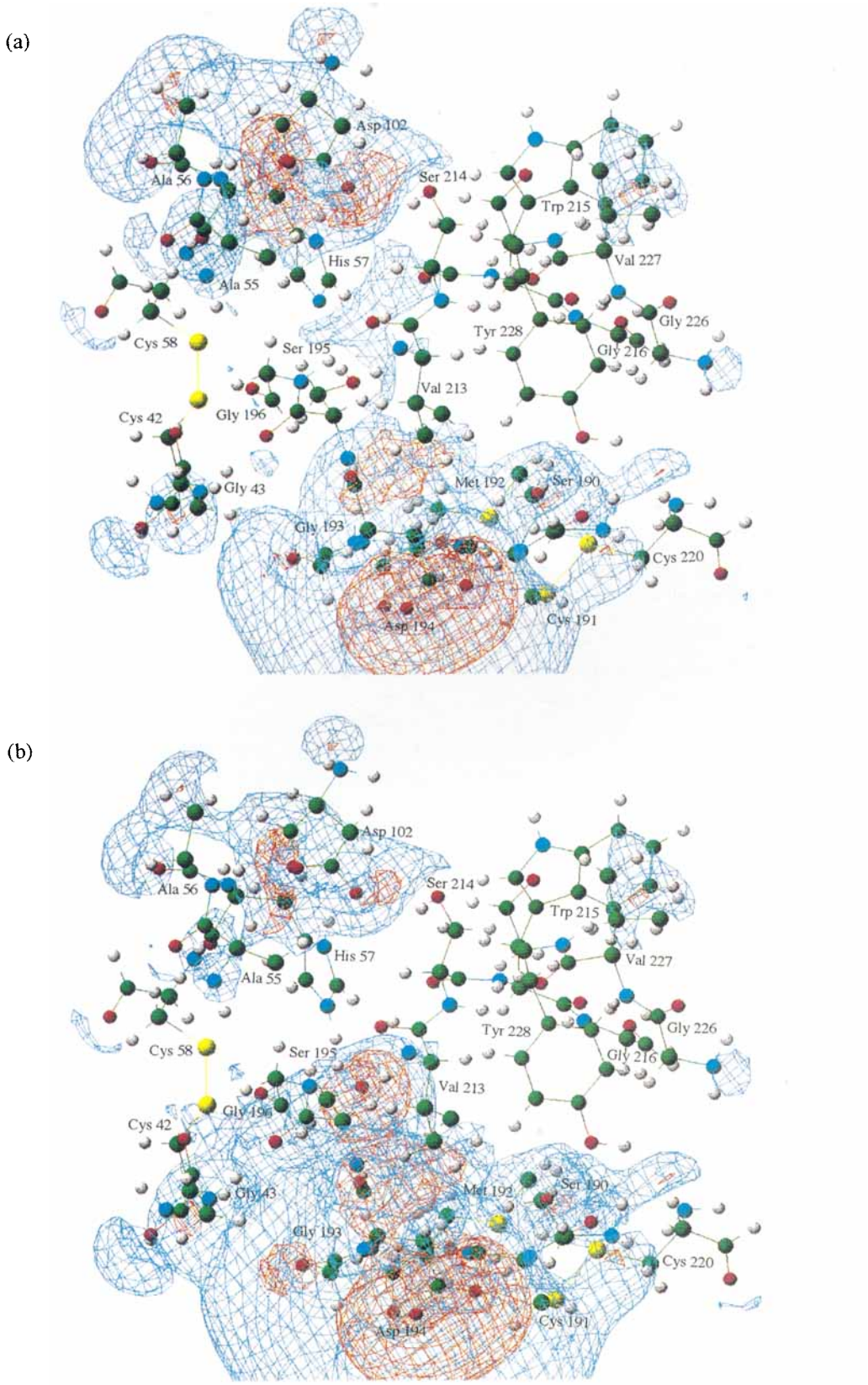


FIGURE 13. 3D FUNC MEP map of the active site of α-chymotrypsin: (a) before proton transfer, and (b) after proton transfer. The active site is composed of 22 amino acids (see text). Isosurfaces were taken at the -80 kcal mol^{-1} (red) and -30 kcal mol^{-1} (blue) levels.

TABLE VIII.
Computation Time Required for the MEP Calculation.[a]

Molecule	STO-5G[b]	FUNC[c]
Glycine	115 + 3039	9
Alanine	252 + 4622	11
Serine	365 + 5956	13
Cysteine	461 + 7144	13
Valine	828 + 8782	15
Aspartate	723 + 9129	16
Asparagine	803 + 9484	17
Methionine	1189 + 12154	16
Histidine	1456 + 13829	21
Tyrosine	2282 + 19501	24
Tryptophan	4533 + 25555	29
Alanylalanine	2098 + 15156	21
Active site of α-chymotrypsin (22 amino acids)	(56000000 + 2500000)[d]	283

[a]All amino acids are in the neutral form. Time is in seconds. The calculation was carried out on the HP 755 workstation.
[b]Time required for SCF-MO calculation (first) and the MEP calculation for 132,651 (51 × 51 × 51) points (second).
[c]The MEP calculation by the functions for 132,651 (51 × 51 × 51) points.
[d]Estimated value.

magnitude of potentials around the nucleophilic sites, and the change in MEP caused by the conformational change were well reproduced by the empirical functions. It is shown that the π electron population of small model molecules can be used successfully for the corresponding fragments in larger molecules and that calculation of the π electron population of a large molecule is not required in the present procedure. The present method can be used successfully for the semiquantitative evaluation of MEPs for large biological molecules and the MEP change caused by the conformational change and/or various types of interactions with very short computation time.

References

1. R. Bonaccorsi, E. Scrocco, and J. Tomasi, *J. Chem. Phys.*, **52**, 5270 (1970).
2. P. Politzer and D. G. Truhlar, *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum, New York (1981).
3. (a) E. Scrocco and J. Tomasi, *Adv. Quantum Chem.*, **11**, 115 (1978); (b) G. G. Hall, *J. Chem. Soc., Faraday Trans. 2*, **85**, 251 (1989).
4. (a) P. Politzer, P. R. Laurence, L. Abrahmsen, B. A. Zilles, and P. Sjoberg, *Chem. Phys. Lett.*, **111**, 75 (1984); (b) P. Politzer and L. Abrahmsen, *J. Am. Chem. Soc.*, **106**, 855 (1984); (c) J. S. Murray, B. A. Zilles, K. Jayasuria, and P. Politzer, *J. Am. Chem. Soc.*, **108**, 915 (1986); (d) P. Politzer, P. Lane, K. Jayasuria, and L. N. Domelsmith, *J. Am. Chem. Soc.*, **109**, 1899 (1987); (e) J. S. Murray and P. Politzer, *J. Mol. Struct. (Theochem)*, **163**, 111 (1988); (f) J. S. Murray and P. Politzer, *J. Mol. Struct. (Theochem)*, **180**, 161 (1988); (g) J. S. Murray and P. Politzer, *J. Mol. Struct. (Theochem)*, **187**, 95 (1989); (h) O. Kikuchi, *J. Mol. Struct. (Theochem)*, **138**, 121 (1986).
5. H. Weinstein, R. Osman, J. P. Green, and S. Topiol, in ref. 2, p. 309.
6. (a) J. S. Murray and P. Politzer, *Theor. Chim. Acta.*, **72**, 507 (1987); (b) J. S. Murray, P. Evans, and P. Politzer, *Int. J. Quantum Chem.*, **37**, 271 (1990).
7. G. Pepe, D. Siri, and J. P. Reboul, *J. Mol. Struct. (Theochem)*, **256**, 175 (1992).
8. (a) A. K. Bhattacharjee and S. N. Bose, *Indian J. Chem.*, **30b**, 991 (1991); (b) B. Testa, N. Van de Waterbeemd, and P.-A. Carrupt, *J. Mol. Struct. (Theochem)*, **134**, 351 (1986); (c) S. W. Guha, D. Majumdar, and A. K. Bhattacharjee, *J. Mol. Struct. (Theochem)*, **256**, 61 (1992).
9. (a) R. P. Sheridan and L. C. Allen, *J. Am. Chem. Soc.*, **103**, 1544 (1981); (b) G. Naray-Szabo, *Int. J. Quantum Chem.*, **22**, 575 (1982).
10. G. Naray-Szabo, *Int. J. Quantum Chem.*, **23**, 723 (1983).
11. J. L.-Brasseur, G. Dive, D. Dehareng, and J.-M. Ghuysen, *J. Theor. Biol.*, **145**, 183 (1990).
12. G. Naray-Szabo, *Int. J. Quant. Chem.*, **16**, 265 (1979).
13. (a) F. A. Momany, *J. Phys. Chem.*, **82**, 592 (1978); (b) U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **5**, 129 (1984); (c) L. E. Cirlian and M. M. Francl, *J. Comput. Chem.*, **8**, 894 (1987); (d) B. Wang and G. P. Ford, *J. Comput. Chem.*, **15**, 200 (1994).
14. F. Manaut, F. Sanz, J. Jose, and M. Milesi, *J. Comp.-Aided Mol. Des.*, **5**, 371 (1991).
15. (a) J. S. Murray and P. Politzer, *J. Org. Chem.*, **56**, 6715 (1991); (b) J. S. Murray and P. Politzer, *J. Org. Chem.*, **56**, 3735 (1991); (c) P. Politzer and W. L. Hedges, *J. Mol. Struct. (Theochem)*, **262**, 155 (1992).
16. (a) P. Politzer, R. G. Parr, and D. R. Murphy, *J. Chem. Phys.*, **79**, 3859 (1990); (b) J. S. Murray, T. Brinck, M. E. Grice, and P. Politzer, *J. Mol. Struct. (Theochem)*, **256**, 29 (1992); (c) M. Haeberlein, J. S. Murray, T. Brinck, and P. Politzer, *Can. J. Chem.*, **70**, 2209 (1992); (d) J.S. Murray, P. Lane, T. Brink, K. Paulsen, M. E. Grice, and P. Politzer, *J. Phys. Chem.*, **97**, 9369 (1993).
17. (a) S. R. Cox and D. E. Williams, *J. Comput. Chem.*, **2**, 304 (1981); (b) H. Kubodera, S. Nakagawa, and H. Umeyama, *Chem. Pharm. Bull.*, **35**, 1673 (1987).
18. (a) R. Bonaccorsi, E. Scrocco, and J. Tomasi, *J. Am. Chem. Soc.*, **98**, 4049 (1976); (b) R. Bonaccorsi, E. Scrocco, and J. Tomasi, *J. Am. Chem. Soc.*, **99**, 4546 (1977).
19. J. S. Murray, J. M. Seminario, M. C. Concha, and P. Politzer, *Int. J. Quantum Chem.*, **44**, 113 (1992).
20. O. Kikuchi, K. Horikoshi, and O. Takahashi, *J. Mol. Struct. (Theochem)*, **256**, 47 (1992).
21. O. Kikuchi, H. Nakajima, K. Horikoshi, and O. Takahashi, *J. Mol. Struct. (Theochem)*, **285**, 57 (1993).
22. D. M. Blow, *Acc. Chem. Res.*, **9**, 145 (1976).
23. H. Tsukada and D. M. Blow, *J. Mol. Biol.*, **184**, 703 (1985).
24. A. Warshel, G. Naray-Szabo, F. Sussman, and J.-K. Hwang, *Biochemistry*, **28**, 3629 (1989).